It is common practice in medical research to investigate the relation between two physiological variables by obtaining pairs of measurements over a range of values and then plotting one variable against the other. The nature of the relation may be examined using a technique such as regression analysis, and the strength of the relation assessed by calculating the correlation coefficient. It is often assumed that when a correlation is found between the variables they are physiologically related in some way. In some situations this may not be correct.

Two variables, each of which is dependent on a third variable which may not be known, can correlate even if not causally linked. For example, unrelated variables that both change with time will correlate. Another cause of spurious correlation is illustrated in figure 1. A random number table was used to generate 50 pairs of numbers, \( x(n) \) and \( y(n) \), between 1 and 10. When these numbers are plotted against one another a random scatter of points is produced and, as expected, the correlation coefficient is 0 (fig. 1A). If the same pairs of numbers are used to calculate a third variable \( x \cdot y(n) \), the product of each data pair, \( x(n) \) re-plotted against \( x \cdot y(n) \) is no longer random. Instead, there is a linear relation with a correlation coefficient of 0.7 and a slope that is statistically different from 0 (fig. 1B). This surprising result occurs because the values of \( x(n) \) are present in both variables (\( x(n) \) and \( x \cdot y(n) \)), which are then linked mathematically. A similar result occurs when variables are derived by addition, subtraction, or division. This is mathematical coupling.

There are many situations in medical research, and in the clinical management of patients, when physiological variables that are difficult to determine directly are calculated from other more easily measured variables. The formulas required are derived from basic physiological principles. Examples in anaesthesia include the deadspace ratio derived from arterial and mixed expired carbon dioxide partial pressures using the Bohr equation, and systemic vascular resistance derived from cardiac output and mean arterial pressure. If the relation is investigated between a derived variable and another variable that is either a component of the derived variable or shares a component with it, the relation may be affected by mathematical coupling. As a result, part or all of the relation may be spurious. Mathematical coupling can occur if paired values for physiological deadspace ratio and arterial carbon dioxide tension, or cardiac output and systemic vascular resistance, are plotted. The confounding effects of mathematical coupling were first pointed out by Archie in 1980 but despite his description, this phenomenon has continued to confuse medical research and is often unrecognized or poorly understood.

A fundamental difference between pairs of random numbers and pairs of physiological data is that the latter may be truly physiologically linked. For example, if paired measurements of cardiac output and arteriovenous oxygen content difference are plotted against one another, a negative correlation might be expected because, under conditions of constant oxygen consumption, increased oxygen delivery to tissues results in decreased oxygen extraction. The product of cardiac output and oxygen difference can be used to calculate oxygen consumption using the Fick principle. In many studies, the relation between oxygen delivery and oxygen consumption has been examined using this method, particularly in the critically ill. Mathematical coupling affects this relation because the cardiac output and arterial oxygen content are used in the calculation of both variables. In many studies in critically ill patients, correlation between oxygen delivery and consumption was thought to indicate an oxygen supply dependency. In these studies it was not clear to what extent this finding was an artefact attributable to mathematical coupling, and this is a subject of considerable controversy. Several recent studies have shown how mathematical coupling operates in oxygen kinetics. A consideration of these, and other relevant work, allow the formulation of guidelines that are applicable to situations in clinical practice or in medical research in which mathematical coupling may occur.

Mathematical coupling is most likely to bias the true relation between two variables which share components if the relation is distorted by non-physiological factors, and then plotted values behave increasingly like the randomly distributed numbers of figure 1. Several mechanisms cause this tendency to randomness in physiological studies. First, if the accuracy of measurement is low, random errors increase the spread of the data, and if the variables have common components they will share the measurement errors associated with those components. Second, if several data from many individuals are combined, the true physiological relation present in an individual patient may be obscured. In either situation, mathematical coupling may cause an artificial correlation. These two processes will be considered in more detail.

Any physiological measurement is subject to measurement error. When the relation between two variables is plotted, large measurement errors increase the scatter of the data making them more random. When these variables have shared components, mathematical coupling of the random errors in both variables can produce a correlation obscuring or changing the true physiological relation. This is most likely when a physiological variable is derived from a number of separate measurements, which makes it almost inevitable that its accuracy is low. For example, arteriovenous oxygen difference is calculated from measurements of haemoglobin concentration, oxygen saturation and oxygen tension in both arterial and mixed venous blood. The errors in these measurements may propagate to make the accuracy of...
the final oxygen difference calculation low. In medical studies, we often look at changes in variables in response to a treatment. The relevance of a change in a variable cannot be assessed if the accuracy of measurements is not known or is not reported, because the change could simply represent random variation. When used appropriately, statistical methods allow for the randomness in a measurement resulting from differences between individuals, or from random errors in the actual measurements. However, even appropriately used statistical tests cannot detect or compensate for the spurious relation produced by mathematical coupling.

The influence of measurement errors on a relation is reduced by increasing the number of points plotted. A common problem in studies liable to bias by mathematical coupling is that there are too few data pairs over too narrow a range of values. This was frequently so in early studies of the oxygen supply-demand relation in critically ill patients: in many studies there were only two data pairs, one before and another after an intervention that itself induced only small alterations in oxygen delivery. Bias caused by mathematical coupling is reduced by studying large changes in the variables of interest and plotting a wide range of values. Stratton and colleagues proposed a statistical method to correct for the effect of mathematical coupling error, which relied on estimating the accuracy of measurements of shared components, and the reliability of any changes observed in them.

Pooling data from different subjects increases the chance of a spurious correlation when there are shared variables.6 This is particularly likely if data from many patients are combined and then analysed, instead of examining each patient individually, or using a single summary statistic from each patient. Pooled data, are more widely scattered than points plotted from a single subject because of interindividual physiological differences. As the scatter widens, it behaves more like the data in figure 1 and has a greater tendency to bias by mathematical coupling. The importance of this source of error in oxygen kinetic studies was shown by Phang and coworkers, who found that there was no relation between oxygen delivery and consumption in most individual patients, but when data from all patients were pooled there was a positive correlation between the variables.

In the critical care literature, there is no doubt that mathematical coupling has confused the interpretation of oxygen kinetic studies in which the Fick method was used to calculate oxygen consumption. Recent studies in which analysis of respiratory gases was used to measure oxygen consumption directly, which avoids mathematical coupling altogether, have found little or no evidence of a dependent relation between oxygen delivery and consumption.2 3 7 8 It is clear from these studies that when oxygen consumption is calculated by the Fick method, mathematical coupling causes a positive correlation between oxygen delivery and consumption that is either entirely artefactual, or much stronger than it really is.

The use of derived physiological variables in medical research and in clinical practice is widespread and sometimes unavoidable. Inevitably there are occasions when mathematical coupling will influence the relation between two variables. We suggest some guidelines drawn from oxygen kinetic studies which are applicable to other areas of medical research (table 1). A greater awareness and understanding of the pitfalls of studying the relation between variables with shared components should reduce the risk of study. If the positive end-point means that more positive than negative errors are included in the analysis, these errors apply to both variables under investigation. This can result in a positive relation caused entirely by mathematical coupling. In this situation, a correctly applied correlation test or regression analysis supports the relation, but the relation is an artefact.
erroneous conclusions in future studies, and avoid repetition of the confusion and controversy that has dogged oxygen kinetics.

T. S. WALSHE A. LEE
Department of Anaesthetics
Royal Infirmary of Edinburgh
Lauriston Place
Edinburgh EH3 9YW

References